Author Names & Affiliations

- Louise Kellogg - Computational Infrastructure for Geodynamics, University of California, Davis
- Catherine Cooper - Washington State University
- Omar Ghattas - University of Texas at Austin
- Frederik Simons - Princeton University

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Geophysics, Geodynamics, Computational Science & Engineering, Applied Math

Title of Submission

Advanced Cyberinfrastructure to Support Research in Computational Geodynamics

Abstract (maximum ~200 words).

Advanced cyberinfrastructure, including advanced modeling software and high-performance computing, has become an essential tool for advancing geophysics research. We outline the needs from our diverse perspective as geophysicists, computational scientists, and applied mathematicians. The scientific research challenges range in scale from modeling seismic rupture and seismic wave propagation occurring over seconds to hours, to mountain building and processes coupling Earth’s surface and interior, to the grand computational challenge of modeling the origins and evolution of Earth’s magnetic field over scales from days to billions of years and have profound and immediate implications for life safety and protection of human infrastructure. To advance these science questions involves advanced cyberinfrastructure including: sustained interdisciplinary collaborations between geoscientists, applied mathematicians, and computational scientists to define the scientific problems and develop the underlying mathematics and algorithms, development of high quality open-source scientific software, training across disciplines to enable use of advanced cyberinfrastructure, the organizational infrastructure to sustain a culture of interdisciplinary collaboration in advanced cyberinfrastrucuture, and access to computing resources ranging from mid-scale to leadership-class computing.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Cyberinfrastructure, including advanced modeling software and high-performance computing, has become an essential tool for advancing geophysics research. The scientific research challenges requiring advanced computational methods range in scale from modeling seismic rupture and seismic wave propagation occurring over seconds to hours, to mountain building and processes coupling Earth's surface and interior, to the grand computational challenge of modeling the origins and evolution of Earth's magnetic field over scales from days to billions of years. Research problems involve interacting systems such as the coupling of fluid flow, elastic deformation, and brittle failure in the Earth's crust associated with induced seismicity and volcanic eruptions. Many of these scientific problems, such as forecasting the occurrence of large, complex multi-fault ruptures and the characteristics of the resulting ground motions from earthquakes; understanding the origins of tsunami, the behavior of volcanos before, during, and after eruptions; the connections between the solid Earth and sea level change; and the variability of the Earth's magnetic field, have profound and immediate implications for life safety and protection of human infrastructure.

A major element connecting these scientific challenges is high-quality, specialized open-source scientific software for modeling the complex physics involved. These research questions span geophysics and computational science; progress requires integrated collaboration among geoscientists, mathematicians, and computer scientists. This collaboration becomes even more critical as the community continues to push the limits of existing modelling efforts as well as initiates efforts to integrate and couple software across subdisciplines (e.g., linking surface processes to interior dynamics or coupling geodynamo to mantle convection simulations).

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Advancing science in many areas requires close collaboration between domain (e.g., earth) scientists and computational scientists. However, only a small subset of domain scientists have a sufficient background to fully engage and collaborate with computational scientists. Scientists who lack computational science skills face significant obstacles in adapting/extending state-of-the-art codes to their needs. As a result, domain scientists need more training in numerical techniques and modern software development practices (repositories, testing, modularity, and extensibility). An additional benefit to meeting this need: teaching domain (e.g. geo- or environmental) scientists how to code is an excellent vehicle for bringing people from diverse or disadvantaged backgrounds together and into computationally-based STEM while raising their level of numeracy. Similarly, computational scientists are rarely trained in the methods used by the scientific domains to pose scientific questions, collect and work with observational data, or develop models; providing them with opportunities to work on

As scientific challenges become greater, developing codes to address these challenges requires more resources and reaches beyond the capabilities of individual investigators and even small groups. The scientific community needs open-source codes that can be adapted and extended to address the needs of specific studies, version control, documentation manuals and training resources and the data-exchange formats and mechanisms that allow for the efficient sharing of information. Community efforts such as the Computational Infrastructure for Geodynamics represent a significant investment in such codes and have been proven to advance the science but are limited in scope by the investment available currently.

Fostering human collaboration networks across the physical and computational sciences requires more than providing physical infrastructure or online communities. It additionally requires establishing familiarity and levels of trust, incentives for reaching-out into collaboration, and the apportioning of credit (e.g. through explicit citation of authored code rather than only recognizing scientific research papers that use scientific software).

Once high-quality software is developed, the associated cyberinfrastructure is required to use it, including high-performance computing facilities, long-term storage and access to large or complex datasets and models, and advanced tools for visualizing data and analyzing results.

Continued commitment to endeavors, such as XSEDE, that provide no-cost access to infrastructure and computational support is integral for the success for the research endeavors of the scientific community. While many universities have centralized computational resources, these are often over-subscribed, under-supported and require buy-in from individual users and/or research groups. Dedicated, no-cost computational time as well as access to computational scientists and system administrators provides researchers with the capacity to sustain research projects, the freedom to explore new directions and the ability to collaborate with other computational researchers on other campuses. Some very large simulations in seismology and geodynamics now require leadership class computing; currently resources for this type of model is limited to a small number of collaborations with national facilities.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

Open-source codes need to be designed for usability by a broad audience, have documentation that describes the scientific background, implementation, and have examples of scientific interest. The authors must receive credit for their contributions to scientific software as they would for more traditionally recognized intellectual contributions to research. Training and workforce development are crucial and to enable a diverse workforce to use advanced cyberinfrastructure requires a multipronged effort involving outreach and multidiscipline collaboration. Communities should leverage technologies for online training as well as building the community through in-person training, workshops, hackathons, and collaborations. The role of NSF as a funding agency should include supporting explicit cross-cutting efforts such as the now defunct, but very successful, “Collaborations in Mathematical Geosciences”.

Support is critically needed for interdisciplinary research at the interfaces of geodynamics, continuum mechanics, applied mathematics, statistics, computer science, and HPC. During the 1990s and 2000s, NSF regularly stood up large cross-cutting programs in computational science (e.g. Grand Challenges, National Challenges, KDI, ITR, PetaApps, CDI). Not surprisingly, this was a period in which major strides were made in addressing challenges in the computational sciences, ones that could not be addressed from the vantage of purely disciplinary programs, and instead required the combined perspective of multiple disciplines (HPC, math, statistics, computer science, and domain science) as represented in multidisciplinary teams. But particularly in the field of computational geodynamics, many challenges remain, and new ones have risen, driven by new advances in computing, models, and data. In response to the need for interdisciplinary research and education in computational science, numerous educational programs, institutes, research centers, and interdisciplinary research groups in this field were created in the past decade at universities across the country. Despite this, NSF’s support for cross-cutting research in computational science in general and applications to the scientific domains in particular has diminished in the past half dozen years, having been replaced largely by data science programs. Yet these data science programs do not address the many outstanding challenges in computational geodynamics, in which data is interpreted through the lens of geophysical models. ACI admirably supports software development for computational science through its SSI program, but this crucial activity is just the last step in a pipeline that include development of data structures, parallel algorithms, meshing, discretizations, linear and nonlinear solvers, and inversion methods. No existing program supports research in these areas that is driven by the needs of specific complex geophysical problems. Software is of course important, but the software is only as good as the upstream models and numerics and algorithms.

In addressing all of the above, it is important to keep in mind that advanced cyberinfrastructure is not just the domain of computer science: developing and sustaining cross-cutting collaborations between geoscientists, mathematicians, computational and statistical (data) scientists is now needed more than ever.